

1 **In the Claims**

2  
3 1. (Currently Amended) A processor-readable medium comprising  
4 processor-executable instructions for personalizing karaoke, the processor-  
5 executable instructions comprising instructions for:

6 segmenting visual content to produce a plurality of sub-shots;

7 segmenting music to produce a plurality of music sub-clips, wherein the  
8 segmenting establishes boundaries between the music sub-clips at beat positions  
9 within the music; and

10 displaying at least some of the plurality of sub-shots as a background to  
11 lyrics associated with the plurality of music sub-clips.

12  
13 2. (Original) The processor-readable medium as recited in claim 1,  
14 additionally comprising instructions for:

15 shortening some of the plurality of sub-shots to a length of a corresponding  
16 music sub-clip from within the plurality of music sub-clips.

17  
18 3. (Original) The processor-readable medium as recited in claim 1,  
19 wherein segmenting the visual content comprises instructions for:

20 dividing a shot into two sub-shots at a maximum peak of a frame difference  
21 curve; and

22 repeating the dividing to result in sub-shots shorter than a maximum sub-  
23 shot length.

1           4.     (Original) The processor-readable medium as recited in claim 1,  
2 additionally comprising instructions for:

3                 filtering the plurality of sub-shots according to importance; and

4                 filtering the plurality of sub-shots according to quality.

5  
6           5.     (Currently Amended) The processor-readable medium as recited in  
7 claim 4, wherein filtering the plurality of sub-shots according to quality comprises  
8 instructions for:

9                 examining color entropy within each of the plurality of sub-shots for  
10 indications of diffusion of color; and

11                 if color entropy is low, analyzing each of the plurality of sub-shots to detect  
12 motion more ~~that~~ than a threshold indicating interest and less than a threshold  
13 indicating low camera and/or object movement;

14                 selecting sub-shots having acceptable motion and/or color entropy scores.

15  
16           6.     (Original) The processor-readable medium as recited in claim 4,  
17 wherein filtering the plurality of sub-shots according to importance comprises  
18 instructions for:

19                 evaluating frames within a sub-shot according to attention indices; and

20                 averaging the attention indices for the frames to determine if the sub-shot  
21 should be included or excluded.

1           7.     (Original) The processor-readable medium as recited in claim 4,  
2 wherein filtering the sub-shots according to importance comprises instructions for:  
3           analyzing for camera motion, for object motion and for specific objects  
4 within the sub-shots;  
5           filtering the sub-shots according to the analysis.

6  
7           8.     (Original) The processor-readable medium as recited in claim 1,  
8 wherein the instructions for segmenting visual content segment video.

9  
10          9.     (Original) The processor-readable medium as recited in claim 8,  
11 additionally comprising instructions for:  
12           selecting important sub-shots from within the plurality of sub-shots; and  
13           selecting sub-shots such that they are uniformly distributed within the  
14 video.

15  
16          10.    (Original) The processor-readable medium as recited in claim 9,  
17 wherein selecting important sub-shots comprises instructions for:  
18           evaluating color entropy, camera motion, object motion and object  
19 detection; and  
20           selecting the important sub-shots based on the evaluation.

11. (Original) The processor-readable medium as recited in claim 9,  
wherein selecting uniformly distributed sub-shots comprises instructions for:

evaluating normalized entropy of the sub-shots along a time line of video  
from which the sub-shots were obtained.

12. (Original) The processor-readable medium as recited in claim 1,  
wherein the instructions for segmenting visual content includes instructions for  
assigning photographs to be sub-shots.

13. (Original) The processor-readable medium as recited in claim 12,  
wherein the instructions for assigning photographs includes instructions for:

rejecting photographs having problems with quality; and

rejecting photographs within a group of very similar photographs wherein a  
photo within the group has been selected.

14. (Original) The processor-readable medium as recited in claim 12,  
wherein the instructions for assigning photographs includes instructions for:

converting at least one of the photographs to video.

15. (Original) The processor-readable medium as recited in claim 1,  
wherein the visual content comprises home video and photographs in digital  
formats.

16. (Cancel)

17. (Original) The processor-readable medium as recited in claim 1, wherein segmenting music into the plurality of music sub-clips comprises instructions for bounding music sub-clip length according to:

minimum length =  $\min\{\max\{2 * \text{tempo}, 2\}, 4\}$  and

maximum length = minimum + 2.

18. (Original) The processor-readable medium as recited in claim 1, wherein segmenting the music comprises instructions for:

establishing music sub-clips' length within a range of 3 to 5 seconds.

19. (Original) The processor-readable medium as recited in claim 18, wherein segmenting the music comprises instructions for:

establishing boundaries for the music sub-clips at sentence breaks.

20. (Original) The processor-readable medium as recited in claim 1, additionally comprising instructions for:

obtaining the lyrics from a file; and

coordinating delivery of the lyrics with the music using timing information contained within the file.

21. (Original) A processor-readable medium as recited in claim 20, wherein obtaining the lyrics comprises instructions for sending the file over a network to a karaoke device as a part of a pay-for-play service.

22. (Original) The processor-readable medium as recited in claim 1, additionally comprising instructions for:

querying a database of songs by humming a portion of a desired song; and

selecting the desired song from among a number of possibilities suggested by an interface to the database.

23. (Currently Amended) A processor-readable medium comprising processor-executable instructions for ~~providing-integrating~~ lyrics, ~~for integration with-music and video content~~ suitable for karaoke, the processor-executable instructions comprising instructions for:

receiving a request for a file associated with a specified song, wherein the file:

~~associates each syllable contained within the lyrics with timing values; and~~

~~associates each sentence contained within the~~ comprises, music, ~~lyrics, and with timing values associated with the lyrics; and~~

fulfilling the request for the file by sending the file associated with the specified ~~song-~~song;

segmenting the music to produce a plurality of music sub-clips, wherein the segmenting establishes boundaries between the music sub-clips at beat positions within the music;

segmenting visual content to produce a plurality of sub-shots of a length corresponding music sub-clips from the plurality of music sub-clips; and

1        outputting the plurality of music sub-clips together with corresponding sub-  
2        shots of visual content, wherein the visual content is configured as a background  
3        to the lyrics associated with the music sub-clips.

4  
5        24. (Original) A processor-readable medium as recited in claim 23,  
6        wherein obtaining the lyrics comprises instructions for sending the file over a  
7        network to a karaoke device.

8  
9        25. (Currently Amended) A personalized karaoke device, comprising:  
10        a music analyzer configured to segment music to produce a plurality of  
11        music sub-clips, wherein the segmenting establishes boundaries between the music  
12        sub-clips at beat positions within the music of a song~~create music sub-clips of~~  
13        ~~varying lengths according to a song;~~

14        a visual content analyzer configured to define and select visual content sub-  
15        shots;

16        a lyric formatter configured to time delivery of syllables of lyrics of the  
17        song; and

18        a composer configured to assemble the music-sub clips with the visual  
19        content sub-shots, and configured to adjust length of the sub-shots to correspond  
20        to the music sub-clips, and configured to superimpose the syllables of the lyrics of  
21        the song over the sub-shots.

1           **26.** (Original) The personalized karaoke device of claim 25, wherein the  
2 music analyzer is configured to segment the song with a strong onset between  
3 each of the music sub-clips.  
4

5           **27.** (Original) The personalized karaoke device of claim 25, wherein the  
6 music analyzer is configured to segment the song with a beat between each of the  
7 music sub-clips.  
8

9           **28.** (Original) The personalized karaoke device of claim 25, wherein the  
10 music analyzer is configured to segment the song automatically into sub-clips,  
11 each having a duration that is a function of song tempo.  
12

13           **29.** (Original) The personalized karaoke device of claim 25, wherein the  
14 visual content analyzer is configured to segment video into sub-shots.  
15

16           **30.** (Original) The personalized karaoke device of claim 25, wherein the  
17 visual content analyzer is configured to access folders of home video and  
18 photographs containing content from which the sub-shots are derived.  
19

20           **31.** (Original) The personalized karaoke device of claim 25, wherein the  
21 visual content analyzer is configured to assemble still photographs, each of which  
22 is a sub-shot.  
23  
24  
25



1           **32.** (Original) The personalized karaoke device of claim 25, wherein the  
2 visual content analyzer is configured to select from among sub-shots according to  
3 ranked importance, wherein importance is gauged by detection of color entropy,  
4 detection of object motion within the sub-shot, detection of camera motion during  
5 the sub-shot, and/or detection of a face within the sub-shot.

6  
7           **33.** (Original) The personalized karaoke device of claim 25, wherein the  
8 visual content analyzer is configured to filter out sub-shots having low image  
9 quality as measured by low entropy and low motion intensity.

10  
11           **34.** (Original) The personalized karaoke device of claim 25, wherein the  
12 visual content analyzer is configured to select sub-shots of greater importance  
13 consistent with creating a uniform distribution of the sub-shots over a runtime of a  
14 source video.

15  
16           **35.** (Original) The personalized karaoke device of claim 25, wherein the  
17 visual content analyzer is configured to reject photographs of low quality by  
18 detecting over and under exposure, overly homogenous images and blurred  
19 images.

20  
21           **36.** (Original) The personalized karaoke device of claim 25, wherein the  
22 visual content analyzer is configured to organize photographs by date of exposure  
23 and by scene, thereby obtaining photographs having a relationship.

37. (Currently Amended) The personalized karaoke device of ~~claim~~  
37~~claim~~ 36, wherein the visual content analyzer is configured to reject  
photographs which are members within a group of very similar photographs,  
wherein one of the group has already been selected.

38. (Original) The personalized karaoke device of claim 25, wherein the  
visual content analyzer is configured to:

detect an attention area within a photograph; and

create a photo to video sub-shot based on the attention area, wherein the  
video includes panning and/or zooming.

39. (Original) The personalized karaoke device of claim 25, wherein the  
lyric formatter is configured to consume a file detailing timing of each syllable  
and each sentence of the lyrics.

40. (Currently Amended) An apparatus, comprising:  
means for creating music sub-clips by segmenting the music to define  
boundaries between the music sub-clips at beat positions of varying lengths  
according to within a song;

means for defining and selecting visual content sub-shots;

means for timing delivery of syllables of lyrics of the song; and

means for assembling the music sub-clips with the visual content sub-shots,  
and to adjust length of the sub-shots to correspond to length of the music sub-  
clips, and to superimpose the syllables of the lyrics of the song over the sub-shots.

1  
2       **41.**   (Original) The apparatus of claim 40, wherein the means for defining  
3 and selecting visual content sub-shots is a video analyzer configured to segment  
4 video into sub-shots.

5  
6       **42.**   (Original) The apparatus of claim 40, wherein the means for defining  
7 and selecting visual content sub-shots is a video analyzer configured to access  
8 folders of home video and photographs containing content from which the sub-  
9 shots are derived.

10  
11       **43.**   (Original) The apparatus of claim 40, wherein the means for defining  
12 and selecting visual content sub-shots is a video analyzer configured for:  
13       detecting an attention area within a photograph; and  
14       creating a photo to video sub-shot based on the attention area, wherein the  
15 video includes panning and zooming.

16  
17       **44.**   (Original) The apparatus of claim 40, wherein the means for timing  
18 delivery of syllables of lyrics of the song is a lyric formatter configured for  
19 consuming a file detailing timing of each syllable and each sentence of the lyrics  
20 and for rendering the lyrics syllable by syllable.